Cloning and Sequence Analysis of the 22-kDa Antigen Genes of *Orientia tsutsugamushi* Strains Kato, TA763, AFSC 7, 18-032460, TH1814, and MAK 119

HONG GE,[a] MIN TONG,[a] ANDREW LI,[b] RAJAN MEHTA,[b] AND WEI-MEI CHING[a]

[a]*Rickettsial Diseases Department, Infectious Diseases Directorate, Naval Medical Research Center, Silver Spring, Maryland 20910, USA*

[b]*Thomas Jefferson High School for Science and Technology Mentorship Program, Alexandria, Virginia 22312, USA*

ABSTRACT: The 22-kDa protein antigen is one of several antigens recognized by sera from scrub typhus patients infected with *Orientia tsutsugamushi*. The 22-kDa protein genes from six *O. tsutsugamushi* strains (Kato, TA763, AFSC 7, 18-032460, TH1814, MAK119) were cloned and their sequences were determined and compared to each other and to the Karp strain sequence listed in GenBank. The sequence alignment revealed that the promoter regions of these seven strains were highly conserved. However, the ORFs exhibited some variation. The phylogenetic analysis of the DNA sequences indicated that among the seven strains assessed, Kato and TA763 were the most closely related, while Karp and TH1814 were the most distantly related. The information gained from this analysis will facilitate our selection of *O. tsutsugamushi* strains from which antigens should be derived to be included in a multivalent vaccine candidate for scrub typhus.

KEYWORDS: *Orientia tsutsugamushi*; scrub typhus; 22-kDa antigen; vaccine

INTRODUCTION

Many different serotypes of *Orientia tsutsugamushi*, the causative agent of scrub typhus, have been documented. The identification of protective antigens, which may be responsible for the different serotypes, is crucial to the understanding of homotypic and heterotypic immunity to *O. tsutsugamushi* and to the development of a scrub typhus vaccine. Western blot analysis of *O. tsutsugamushi* cell lysates with patient sera has identified at least five protein antigens of molecular mass 110, 56, 47, 35, and 22 kDa. Among these antigens, the variable 56-kDa protein antigen is the most

Address for correspondence: Dr. Hong Ge, Rickettsial Diseases Department, Infectious Diseases Directorate, Naval Medical Research Center, Silver Spring, Maryland 20910, USA. Voice: 301-319-7439; fax: 301-319-3015.

geh@nmrc.navy.mil

Disclaimer: This work was supported by Work Unit Number A0311. The views expressed in this article are those of the author and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, nor the U.S. Government.

Ann. N.Y. Acad. Sci. 1063: 231–238 (2005).
doi: 10.1196/annals.1355.036

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2005	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Cloning and Sequence Analysis of the 22-kDa Antigen Genes of Orientia tsutsugamushi Strains Kato, TA763, AFSC 7, 18-032460, TH1814, and MAK 119	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Rickettsial Diseases Department, Infectious Diseases Directorate, Naval Medical Research Center, Silver Spring, Maryland 20910**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) **Naval Medical Research Center 503 Robert Grant Avenue Silver Spring, MD 20910-7500**	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**SAR**	**8**	

abundant and is recognized by almost all scrub typhus patients' sera.[1] The 47-kDa antigen is believed to be the most conserved antigen, while the 56-kDa and 110-kDa protein antigens exhibit considerable variations. The sequences of 22-kDa antigens are not available from strains other than prototype Karp strain.

It has been shown that the 22-kDa protein contains B- and T- cell epitopes and that a recombinant 22- kDa protein could induce antigen-specific proliferation of T-cells from mice immunized with whole cell *O. tsutsugamushi.*[2] Many studies have demonstrated that protective immunity to scrub typhus is led by specific T cells; especially Th1 cells. These results suggest that it may be beneficial to include 22-kDa antigen in a multicomponent vaccine candidate. Therefore, we amplified 22-kDa genes from six strains (Kato, TA763, AFSC 7, 18-032460, TH1814, MAK119) and compared their sequences to each other and to that of the Karp strain. The information obtained in this study from the different strains will facilitate our selection of strains to be included in the development of a multivalent vaccine against scrub typhus.

MATERIALS AND METHODS

Purification of Genomic DNA, PCR Amplification, Cloning and Fast-Screening of the 22-kDa Gene

The genomic DNAs were purified from *O. tsutsugamushi*–infected mouse spleens. Cell pellet was resuspended in ATL lysis buffer (Qiagen, Valencia, CA). After proteinase K digestion and RNase treatment according to manufacturer instructions, DNA was extracted by phenol-chloroform-isoamyl alcohol three times and precipitated with ethanol-sodium acetate. Purified *O. tsutsugamushi* genomic DNAs from six strains (Kato, TA763, AFSC 7, 18-032460, TH1814, MAK 119) were used as templates for PCR. The primers were designed for amplification of the entire 22-kDa gene based on the published sequence of Karp strain (forward: 5′ TAA GTC GAC ATA CTA TAT CAT CTA CAC AGA AGT TG 3′ and reverse: 5′ TAA GGC ATA TGT ATT CTA CTA TAG CTT GGG GT 3′). The PCR yielded a 1,211 bp fragment that was cloned into pPCR-Script Amp SK+ vector and the positive clones were selected with the Epicentre® Colony Fast-Screen™ Kit (Madison, WI).

Sequencing, Sequence Alignment, and Phylogenetic Analysis of the Amplified 22-kDa Gene

The positive clones were sequenced on the ABI 370 analyzer (Foster City, CA). The inserted sequences from three colonies of each strain were determined. The chromatography was manually examined to resolve any discrepancies in sequencing. The confirmed sequences of six strains were aligned for comparison with that of the Karp strain and phylogenetic trees were constructed by CLUSTALW comparison (version 1.82).

RESULTS

Sequence alignment of seven 22-kDa genes (six determined in this study and the published sequence of Karp strain) demonstrated that the promoter regions were

```
Kato            TAAGTCGACATACTATATCATCTACACAGAAGTTGAGTAAAAAATTGCAGCTATGTATTT 60
TA763           TAAGTCGACATACTATATCATCTACACAGAAGTTGAGTAAAAAGCTACAGCTATGTATTT 60
AFSC7           TAAGTCGACATACTATATCATCTACACAGAAGTTGAGTAAAAAACTGCAGCTATGTATTT 60
Karp            TAAGTCGACATACTATATCATCTACACAGAAGTTGAGTAAAAAACTGCAGCTATGTATTT 60
18-032460       TAAGTCGACATACTATATCATCTACACAGAAGTTGAGTAAAAAGCTACAGCTATATATTT 60
TH1814          TAAGTCGACATACTATATCATCTACACAGAAGTTGAGTAAAAAACTGCAGCTATGTATTT 60
MAK119          TAAGTCGACATACTATACCATCTACACAGAAGTTGAGTAAAAAACTGCAGCTATGTATTT 60
                ********************************************  * ******* *****

Kato            CATTAGCAACTTTTATATTACGTTGCATCCTATATGCTTCACGATTAAAAATATAATTCA 120
TA763           CATTAGCAACTTTTATATTACGTTGCATCCTATATGCTTCACGATTAAAAATATAATTCA 120
AFSC7           CATTAGCAACTTTTATATTACGTTGCATCCTATATGCTTCACGATTAAAAATATAATTCA 120
Karp            CATTAGCAACTTTTATATTACGTTGCATCCTATATGCTTCACGATTAAAAATATAATTCA 120
18-032460       CATTAGCAACTTTTATATTACGTTGCATCCTATATGCTTCACGATTAAAAATATAATGCA 120
TH1814          CATTAGCAACTTTTATATTACGTTGCATCCTATATGCTTCACGATTAAAGATATAATTCA 120
MAK119          CATTAGCAACTTTTATATTACGTTGCATCCTATATGCTTCACGATTAAAAATATAATTCA 120
                ************************************************ ******* **

Kato            TAGCCTGTTTTAATA-TTTAGCTAACTAAAAGAAATATTTTGTAAGATAATACAGTATAT 179
TA763           TAGCCTGTTTTAATA-TTTAGCTAACTAAAAGAAATATTTTGTAAGATAATACAGTATAT 179
AFSC7           TAGCCTGTTTTAATA-TTTAGCTAACTAAAAGAAATATTTTGTAAGATAATACAGTATAT 179
Karp            TAGTCTGTTTTAATA-TTTAGCTAACTAAA-GAAATATCCTGCAAGATAATACAGTATGT 178
18-032460       TAGTCTGTTTTAATAATTTAGCTAACTAAA-GAAATATCTTGCAAGATAATACAGTATAT 179
TH1814          TAGCCTGTTTTAATA-TTTAGCTAACTAAAAGAAATATTTTGTAAGATAATATAGTATAT 179
MAK119          TAGCCTGTTTTAATA-TTTAGCTAACTAAAAGAAATATTTTGTAAGATAATACAGTATAT 179
                *** *********** ************** *******  ** ********* ***** *
                                                      -75
Kato            GTATAATCATGATATATGATAGAGGATCAATACAACTGAAACAATATTAGTTATTGATAA 239
TA763           GTATAATCATGATATATGATAGAGGATCAATACAACTGAAACAATATTAGTTATTGATAA 239
AFSC7           GTATAATCATGATATATGATAGAGGATCAATACAACCGAAACAATATTAGTTATTGATAA 239
Karp            GTATAATCATGATATATGATAAATGATCAATATAACTGAAATAATATTAGTTATTGATAA 238
18-032460       GTATAATTATTATATGTGATAGAGTATCAATACAATTGAAATAATATTAGTTATTGATAA 239
TH1814          GTATAATCATGATATATGATAGAGGATCAATACAACTGAAACAATATTAGTTATTGATAA 239
MAK119          GTATAATCATGATATATGATAGAGGATCAATACAACTGAAACAATATTAGTTATTGATAA 239
                ******* ** **** ***** *  ******* **  **** ******************
                                                                  start
Kato            CTGCATAATATTTAAGTATATTTTTAAGTATTACTAATATTATAAGGAATTATCTTTATG 299
TA763           CTGCATAATATTTAAGTATATTTTTAAGTATCACTAATATTATAAGGAATTATCTTTATG 299
AFSC7           CTGCATAATATTTAAGTATATTTTTAAGTATTACTAATATTATAAGGAATTATCTTTATG 299
Karp            CTGCATAATATTTAAGTATATTTTTAAGTATTACTAATATTATAAGGAATTATCTTTATG 298
18-032460       CTGCATAATATTTAAGTATATTTTTAAGTATTACTAATATTATAAGGAATTATCTTTATG 299
TH1814          CTGCATAATATTTAAGTATATTTTTAAGTATTACTAATATTATAAGGAATTATCTTTATG 299
MAK119          CTGCATAATATTAGAGTATATTTTTAAGTATTACTAATATTATAAGGAATTATCTTTATG 299
                ************  ***************** ****************************

Kato            AGTAAAGAAGCAACAGAACTCAAAGAAAATATGAATAAAGCAATGGAGTCTGCATCGCAA 359
TA763           AGTAAAGAAGCAACAGAACTCAAAGAAAATATGAATAAAGCAATGGAGTCTGCATCGCAA 359
AFSC7           AGTAAAGAAGCAACAGAACTCAAAGAAAGTATGGATCAAGCAATGAAGTCTGCATCGCAA 359
Karp            AGTAAAGAAGCAACAGAACTCAAAGAAAGTATGGATCAAGCAATGAAGTCTGCATCGCAA 358
18-032460       AGTAAAGAAGCAACAGAACTCAAAGAAAGTATGGATCAAGCAATGAAGTCTGCATCGCAA 359
TH1814          AGTAAAGAAGCAACAGATCTCAAAGAAAATATGAATAAAGCAATGGAGTCTGCATCCCAA 359
MAK119          AGTAAAGAAGCAACAGATCTCAAAGAAAATATGAATAAAGCAATGGAGTCTGCATCCCAA 359
                ***************** ********** **** ** ******** ********** ***

Kato            AATTCAAGTATTTCAGCTGAACAAAAATTACAACTTGAAAACGAAAAAAAGGAACTGCAA 419
TA763           AATTCAAGTATTTCAGCTGAACAAAAAGCACAACTTGAAAATGAAAAAAAGGAACTGCAA 419
AFSC7           AATTCAAGTATTTCAGCTGAACAAAAATTACAACTTGAAAAAGAAAAAAAGGAACTGCAA 419
Karp            AATTCAAGCATTTCAGCTGAACAAAGATCACAACTTGAAAAAGAAAAAAAGGAACTGCAA 418
18-032460       AATTCGAGTATTTCAGCTGAACAAAAATTACAACTTGAAAACGAAAAAAAGGAACTGCAA 419
TH1814          AGTTCAAGCCTTTCAGCTAAACAAAGATCAGAACTTGAAACTGAAAAAAATGAACTGCAA 419
MAK119          AGTTCAAGCCTTTCAGCTAAACAAAAATCAGAACTTGAAACTGAAAAAAATGAACTGCAA 419
                * *** **  ******** ****** *  * *********  ******** *********
```

FIGURE 1. The alignment of 22-kDa genes from seven *O. tsutsugamishi* strains.

```
Kato            GGACAAATTAGCGATATTACTGGAAAAGATGAACAATCTCAACAAACTACAATGGATAAA 479
TA763           GGCCAAATTAGCGATATTGCTGGAAAAGATGAACAAACCCAACAAACTACAATGGATAAA 479
AFSC7           GGCCAAATTGGCGATATCACTGGGAAAGATGAACAAACCCAACAAACTACAATGAATAAA 479
Karp            GGCCAAATTGGCGATATCACTGGGAAAGATGAACAAACCCAACAAACTACAATGAATAAA 478
18-032460       GGCCAAATTGGCGATATCACTGGGAAAGATGAACAAACCCAACAAACTACAATGAATAAA 479
TH1814          AGCCAAATTAGCGCTATTACTGGAAAAGATAAACAATCTCAAGAAACTACAATGGCTAAA 479
MAK119          AGCCAAATTAGCGCTATTACTGGAAAAGATAAACAATCTCAAGAAACTACAATGGCTAAA 479
                 * ****** *** ***  **** ****** ***** * *** ***********  ****

Kato            TTGAAAGAATGGATGTTAAAGATCAAGGAATTTATCACAAGTAAGGATTTTTCAGAGCTA 539
TA763           TTAAAAGAATGGATGTTAAAGATCAAGGAATTTATCACAAGTAATGATTTTTCAAAGCTA 539
AFSC7           TTAAAAGAATGGATGTTAAAGATCAAGGATTTTCTTACAAGTGATGATTTTTCAAAGCTA 539
Karp            TTAAAAGAATGGATGTTAAAGATCAAGGATTTTCTTATAAGTGATGATTTTTCAAAGCTA 538
18-032460       TTAAAAGAATGGATGTTAAAGATCAAGGATTTTCTTACAAGTGATGATTTTTCAAAGCTA 539
TH1814          TTGAAAGAATGGATGTTAAAGATCAAGGAATTTATTACAAGTCAGGATTTTTCCAATCTA 539
MAK119          TTGAAAGAATGGATGTTAAAGATCAAGGAATTTATTACAAGTCAGGATTTTTCCAATCTA 539
                ** ************************** *** * * **** * ********  * ***

Kato            GTAGATAGCGTAGTAAAATTTGTACAAACTGCAGTCAAAGTTTCCACAGAAATGGCACAA 599
TA763           GTAGATAGCGCAGTGCAATTTGTACAAACAGCAGTCAAAGTTTCCGCAGAAATGGTGCAG 599
AFSC7           GTAGATAGCGCAGTAAAATTTGTACAAACTGCAGTTAAGGTTTCCACAGAAATGATGCAG 599
Karp            GTAGATAGCGCAGTAAAATTTGTACAAACTGCAGTTAAGGTTTCCACAGAAATGATGCAG 598
18-032460       GTAGATAGCGCAGTAAAATTTGTACAAACTGCAGTTAAGGTTTCCACAGAAATGATGCAG 599
TH1814          GTAGATAGTGCAGTAAAATTTGTACAAACTGCAGTTAAGGTCTCTACAGAAATGATGAAG 599
MAK119          GTAGATAGTGCAGTAAAATTTGTACAAACTGCAGTTAAGGTCTCTACAGAAATGATGAAG 599
                ******** * ***  ************** ***** ** ** **  ********    *

Kato            GCCTTTACAGGCATGAAAGATAAAGGAATAATGGGAGTAGCTGCAGGCATACAAACTGTT 659
TA763           GCTTTTACAGGCATGAAAGAGAAAGGAATAATGGGTGTAGCTGCAGGCATAAAAACTGTT 659
AFSC7           GCCTTTACAGGCATGAAAGAGAAAGGAATAATGGGTGTAGCTGCAGGCATACAAACTGTT 659
Karp            GCCTTTACAGGCATGAAAGAGAAAGGAATAATGGGTGTAGCTGCAGGCATACAAACTGTT 658
18-032460       GCTTTTACAGGCATGAAAGAGAAAGGAATAATGGGAGTAGCTGCAGGAATACAAACTGTT 659
TH1814          GCTTTTACAGGCATGAAAGAGAAAGGAATAATGGGTGTAGCTGAAGGCATACAAACTGTT 659
MAK119          GCTTTTACAGGCATGAAAGAGAAAGGAATAATGGGTGTAGCTGAAGGCATACAAACTGTT 659
                ** ***************** ************** ******* *** *** ********

Kato            ACTAGTGGATTCCAAGATATAACCCAAGGTGTGAGCAAGATGGTTGAAGCTGGAGAAGCA 719
TA763           ACTAGTGGATTCCAAGATATAACCCAAGGTGTGAGCAGGATGGTTGAAGCTGGAGAAGCA 719
AFSC7           ACTAGTGGATTTCAAAATATAACACAAGGGGTGAGCAATATGATTGAAGCTGGAGAAGCA 719
Karp            ACTAGTGGATTTCAAAATATAACACAAGGGGTGAGCAATATGATTGAAGCTGGAGAAGCA 718
18-032460       ACTACTGGATTCCAAAGTATAACCCAAGGTGTGAGCAATATGATTGAAGCTGGAGAAGCA 719
TH1814          ACTACTGGATTCCAAAGTATAACCCAAGGTGTGAGCAATATGATTGAAGCTGGAGAAGCA 719
MAK119          ACTACTGGATTCCAAAGTATAACCCAAGGTGTGAGCAATATGATTGAAGCTGGAGAAGCA 719
                **** ****** ***  ****** ***** ******* *** ******************

Kato            GCAGTAAAGCACTTTTCTTCTTCTGGAGATAAAATGGGAAAGAAATTTTTAGGAGCAGAG 779
TA763           GCAGCAAAGCACTTTTCTTCTTCTGGAGATAAAAT---AAAAGAAATTTTAGGAGCAGAG 776
AFSC7           GCAGTAAAACACTTTTCTTCTCCTAGAGATAAAAT---AAAAGAAGCTTTAGGAGCAGAA 776
Karp            GCAGTAAAACACTTTTCTTCTCCTAGAGATAAAAT---AAAAGAAGCTTTAGGAGCAGAA 775
18-032460       GCAGCAAAGCACTTTTCTTCTCCTGGCGATAAAAT---AAAAGAAGCTTTAGGAGCAGAG 776
TH1814          GCAGCAAAGCACTTTTCTTCTCCTGGCGATAAAAT---AAAAGAAGCTTTAGGAGCAGAG 776
MAK119          GCAGCAAAGCACTTTTCTTCTCCTGGCGATAAAAT---AAAAGAAGCTTTAGGAGCAGAG 776
                **** *** ************ ** * ********   ***  **  ************
```

FIGURE 1 — *continued*.

```
Kato            GGATTAGCTAAACTTCAAGCTGCTAGCGCTGGATTACAAAATAATGCTAGTATAGCATCA 839
TA763           GGATTAGCTAAACTTCAAGCTGCTAGCGCTGGATTACAGAGTAATGCTAGTATAGCATCA 836
AFSC7           GGACTAGCTAAACTTCAAGCTGCTAGCGCTGGATTACAAAATAATGCTAGTATAGCATCA 836
Karp            GGATTAGCTAAACTTCAAGCTGCTAGCGCTGGATTACAAAATAATGCTAGTATAGCATCA 835
18-032460       GGACTAGCTAAACTTCAAGCTGCTAGCGCTGGATTACAAAGTAATGCTAGTATAGCATCA 836
TH1814          GGACTAGCTAAAACTTCAGCTGCTAGTGCTGGATTACAAAATAATGCTAGTATAGCATCA 836
MAK119          GGACTAGCTAAACTTCAAGCTGCTAGTGCTGGATTACAAAATAATGCTAGTATAGCATCA 836
                *** ********  *  ********* *********** * ********************

Kato            CCTTCAGCTAGTAGTGTATCAACTCCAACAGCAGCGCGC------ACACAACAGAAAGAC 893
TA763           CCTTCAGCTAGTAGTGTATCAACTCCAACAGCAGCGCGCTCAACAACACCACAGAAAGAC 896
AFSC7           TCTTCAGCTAGT---GTATCAACTCCAACAGCAGCACGC------CCACAACAGAAAGAC 887
Karp            TCTTCAGCTAGT---GTATCAACTCCAACAGCAGCGCGC------ACACAACAGAAAGAC 886
18-032460       TCTTCAGCTAGT---GTATCAACTCCAACAGCAGCACGCTCC---CCACAACAGAAAGAC 890
TH1814          TCTTCAGCTAGT---GTATCAACTCCAACAGCAGCACGT------ACACAACAGAAAGAC 887
MAK119          TCTTCAGCTAGT---GTATCAACTCCAACAGCAGCACGT------ACACAACAGAAAGAC 887
                 ***********   ********************* **        *** **********

                          stop
Kato            ---TCTATAGCAAGATAAAACA-GTATGATAACACTAGAATAGGATAATTTTAAGCGTGT 949
TA763           AGTTCTATAGCAAGATAAAACA-GTATTATAACACTAGGATAGGATAATTTTAAGCGTGT 955
AFSC7           ---TCTATAGCAAGATAAAACAAGTATTCTAACGCTAGAATAGAATAATTTTAAGCGCGT 944
Karp            AGTTCTATAGCAAGATAAAACA-GTATTATAACGCTAGAATAGAATAATTTTAAGCGCGT 945
18-032460       ---TCTATAGCAAGATAAAACA-GTATGATAACACTAGAATAGGATAATTTTAAGCGTGT 946
TH1814          AGTTCTATAGCAAGATAAAACA-GTATGATAACACTAGAATAGGATAATTTTAAGCGTGT 946
MAK119          AGTTCTATAGCAGGATAAA-CA-GTATGATAACACTAGAATAGGATAATTTTAAGCGTGT 945
                   ********* ****** ** ****  **** **** **** ************* **

Kato            TAGCCAAGAAAAAAGAAAAA---AACAGATATAAGAAAAAATTAAGTTTAAAGT-TAAAA 1005
TA763           TTG-------AGAGAAAAAA---CATAGATATAAGAAAAAATTAAGTTTAAAGTCTAAAA 1005
AFSC7           GTTAGCGAAGAAAGAAAAAAAAACATAGATATAAGAAAAAATTAAGTTTAAAGT-TAAAA 1003
Karp            GTTAGCGAAGAAAGAAAAAA--ACATAGATATAAGAAAAAATTAAGTTTAAAGT-TAAAA 1002
18-032460       TT---------GAGAAAAAA---CATAGATATAAGAAAAAATTAAGTTTAAAGT-TAAAA 993
TH1814          TG---------GAGAAAAAA---CATAGATATAAGAAAAAATTAAGTTTAAAGT-TAAAA 993
MAK119          TG---------GAGAAAAAA---CATAGATATAAGAAAAAATTAAGTTTAAAGT-TAAAA 992
                          *  *****    * ***************************** *****

Kato            TAACTTGCAACTCAATTAATCTAGTAGATATGTTTTAATCTTTACGAAGATTAAATAAAT 1065
TA763           TAACTTGCAACTCAATTAATCTAGTAAATATGTTTTAATCTTTACGAAGATTAAATAAAT 1065
AFSC7           TAACTTGCAACTCAATTAATCTAGTAGATATGTTTTAATCTTTACGAAGATTAAATAAAT 1063
Karp            TAACTTGCAACTCAATTAATCTAGTAGATATGTTTTAATCTTTACGAAGATTAAATAAAT 1062
18-032460       TAACTTGCAACTCAATTAATCTAGTAGATATGTTTTAATCTTTACGAAGATTAAATAAAT 1053
TH1814          AAACTTGCAACTCAATTAATCTAGTAAATATGTTTTAATCTTTACGAAGGGAAAATAAAT 1053
MAK119          AAACTTGCAACTCAATTAATCTAGTAAATATGTTTTAATCTTTACGAAGATTAAATAAAT 1052
                 ************************* ***********************   ********

Kato            TAAGTATAAAGGATACTTTTGCTTACCCCAAGCTATAGTAGAATAATACATATGCCTTA 1124
TA763           TAAGTATAAAGGATACTTTTGCTTACCCCAAGCTATAGTAGAATAATACATATGCCTTA 1124
AFSC7           TAAGTATAAAGGATACTTTTGCTTACCCCAAGCTATAGTAGAATAATACATATGCCT-- 1120
Karp            TAAGTATAAAGGATACTTTTGCTTACCCCAAGCTATAGTAGAATAATACATATGCCTTA 1121
18-032460       TAAGTATAAAGGATACTTTTGCTTACCCCAAGCTATAGTAGAATAATACATATGCCTT- 1111
TH1814          TAAGTATAAAGGATACTTTTTCTTACCCCAAGCTATAGTAGAATAATACATAT-CCTTA 1111
MAK119          TAAGTATAAAGGATACTTTTGCTTACCCCAAGCTATAGTGGAATAATACATATGCCTTA 1111
                ******************** ****************** ************* ***
```

FIGURE 1 — *continued*.

highly conserved, while the open reading frames (ORFs) showed sequence variations scattered all over the region (FIG. 1). In the promoter regions, the different strains' upstream sequences between –1 to –75 bp before the ATG start codon were almost identical except for two nucleotides of MAK119 (242, 243) and one nucleotide of TA763 (261). However, there were quite a lot of single nucleotide differences throughout the entire ORF among these seven strains. In addition, there were five triple nucleotide insertions or deletions within the 150 bp region before the stop codon relative to Karp 22-kDa sequence. Using the Karp strain sequence as the reference, percentages of sequence homology observed among these strains if the insertions and deletions were not counted were as follows: AFSC7, 99.46%; 18-032460, 98.13%; TA763, 96.25%; Kato, 96.07%; MAK119, 95.27%; TH1814, 95.18%. The alignment for the deduced protein sequences of seven strains was performed as well (FIG. 2). Although amino acid variations were found scattered throughout the whole ORF, the results indicated that the 22-kDa protein is relatively conserved. Their relationships to Karp strain were similar to those seen in DNA sequence comparisons. The percentages of protein sequence homology among these

```
Kato            MSKEATELKENMNKAMESASQNSSISAEQKLQLENEKKELQGQISDITGKDEQSQQTTMD
TA763           MSKEATELKENMNKAMESASQNSSISAEQKAQLENEKKELQGQISDIAGKDEQTQQTTMD
AFSC7           MSKEATELKESMDQAMKSASQNSSISAEQKLQLEKEKKELQGQIGDITGKDEQTQQTTMN
18-032460       MSKEATELKESMDQAMKSASQNSSISAEQKLQLENEKKELQGQIGDITGKDEQTQQTTMN
TH1814          MSKEATDLKENMNKAMESASQSSSLSAKQRSELETEKNELQSQISAITGKDKQSQETTMA
MAK119          MSKEATDLKENMNKAMESASQSSSLSAKQKSELETEKNELQSQISAITGKDKQSQETTMA
Karp            MSKEATELKESMDQAMKSASQNSSISAEQRSQLEKEKKELQGQIGDITGKDEQTQQTTMN
                ******:***.*::**:****.**:**:*: :**.**:***.**. *:***:*:*:***

Kato            KLKEWMLKIKEFITSKDFSELVDSVVKFVQTAVKVSTEMAQAFTGMKDKGIMGVAAGIQT
TA763           KLKEWMLKIKEFITSNDFSKLVDSAVQFVQTAVKVSAEMVQAFTGMKEKGIMGVAAGIKT
AFSC7           KLKEWMLKIKDFLTSDDFSKLVDSAVKFVQTAVKVSTEMMQAFTGMKEKGIMGVAAGIQT
18-032460       KLKEWMLKIKDFLTSDDFSKLVDSAVKFVQTAVKVSTEMMQAFTGMKEKGIMGVAAGIQT
TH1814          KLKEWMLKIKEFITSQDFSNLVDSAVKFVQTAVKVSTEMMKAFTGMKEKGIMGVAEGIQT
MAK119          KLKEWMLKIKEFITSQDFSNLVDSAVKFVQTAVKVSTEMMKAFTGMKEKGIMGVAEGIQT
Karp            KLKEWMLKIKDFLISDDFSKLVDSAVKFVQTAVKVSTEMMQAFTGMKEKGIMGVAAGIQT
                **********:*: *.***:****.*:**********:** :******:******* **:*

Kato            VTSGFQDITQGVSKMVEAGEAAVKHFSSSGDKMGKKFLGAEGLAKLQAASAGLQNNASIA
TA763           VTSGFQDITQGVSRMVEAGEAAAKHFSSSGDKI-KEILGAEGLAKLQAASAGLQSNASIA
AFSC7           VTSGFQNITQGVSNMIEAGEAAVKHFSSPRDKI-KEALGAEGLAKLQAASAGLQNNASIA
18-032460       VTTGFQSITQGVSNMIEAGEAAAKHFSSPGDKI-KEALGAEGLAKLQAASAGLQSNASIA
TH1814          VTTGFQSITQGVSNMIEAGEAAAKHFSSPGDKI-KEALGAEGLAKTSAASAGLQNNASIA
MAK119          VTTGFQSITQGVSNMIEAGEAAAKHFSSPGDKI-KEALGAEGLAKLQAASAGLQNNASIA
Karp            VTSGFQNITQGVSNMIEAGEAAVKHFSSPRDKI-KEALGAEGLAKLQAASAGLQNNASIA
                **:***.******.*:******.*****. **: *: ******** .*******.*****

Kato            SPSASSVSTPTAAR--TQQKDS-IAR
TA763           SPSASSVSTPTAARSTTPQKDSSIAR
AFSC7           SSSAS-VSTPTAAR--PQQKDS-IAR
18-032460       SSSAS-VSTPTAAR--SPQQKDSIAR
TH1814          SSSAS-VSTPTAAR--TQQKDSSIAR
MAK119          SSSAS-VSTPTAAR--TQQKDSSIAG
Karp            SSSAS-VSTPTAAR--TQQKDSSIAR
                *.*** ********  . *:.. **
```

FIGURE 2. The alignment of amino acids of 22-kDa protein genes from seven *O. tsutsugamishi* strains.

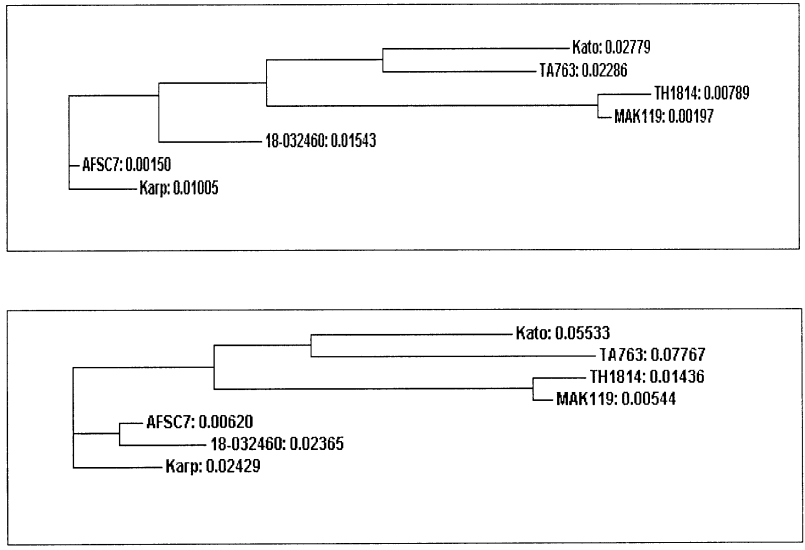

FIGURE 3. The phylogenetic trees for the ORFs of 22-kDa genes (*top*) and protein sequence (*bottom*).

strains compared to Karp strain were 98.02%, 93.07%, 86.14%, 86.14%, 84.65%, and 84.65% in AFSC7, 18-032460, Kato, TA763, TH1814, and MAK119, respectively, if insertions or deletions were excluded.

Phylogenetic trees were constructed using DNA sequences of the ORF (611 bp) and the translated protein sequences of 22-kDa genes by CLUSTALW method (FIG. 3a and b). The lengths of tree lines were proportional to the distances among strains. These distances were also labeled by bootstrap values after the strain names. Based on the DNA sequences of 22-kDa ORF (FIG. 3a), the strains of Kato and TA763 were grouped into the same cluster, while the strains of TH1814 and MAK119, and the strains of AFSC7 and Karp were grouped into two separate clusters. The distance of ORF between strains calculated in this method suggested that the most divergent strains were Karp and TH1814, while Kato and TA763 were the most closely related strains (FIG. 3a). The protein sequence of 22-kDa protein was analyzed by CLUSTALW 1.82 as well (FIG. 3b). Overall, the phylogenetic trees from DNA and protein sequences look similar. But in protein analysis, the 18-032460 instead of Karp was grouped into the same cluster with AFSC7 and the most divergent strains were 18-032460 and TA763.

DISCUSSION

O. tsutsugamushi isolates are antigenically heterogeneous.[3] Therefore, the development of a multivalent vaccine against *O. tsutsugamushi* that provides heterologous

protection has been a major concern. The 22-kDa antigen gene, a possible scrub typhus vaccine component, was first cloned and expressed using the template from Karp strain.[2] However, knowing the 22-kDa sequences from other strains is essential in determining the level of disparity that exists among the various strains of *O. tsutsugamushi*. In this study, we sequenced and analyzed 22-kDa genes from six *O. tsutsugamushi* strains. The results from DNA and protein sequence analysis provided the first sequence comparison and phylogenetic relationship among these *O. tsutsugamushi* strains based on the 22-kDa antigen. The data also extended the earlier findings of the 22-kDa gene from Karp strain and will be helpful in the design and development of scrub typhus vaccines.

The search of the BLAST database for the 22-kDa DNA or protein sequence did not retrieve any sequences with significant homology. There is no conserved domain found on this protein as well. Therefore, the function of this 22-kDa protein remains unknown. However, previous study[2] showed that native 22-kDa protein were reactive with immune serum against homologous strains. The recombinant Karp 22-kDa protein was recognized by polyclonal rabbit antibodies raised against whole-cell lysate. Moreover, the computational analysis of amino acid sequence revealed several potential T-cell epitopes on this protein and this recombinant protein induced a strong proliferation of a T cell line that produced IL-2 and IFN-γ.[2] All this information has strengthened the idea of including the 22-kDa protein as one of the protective antigens in multicomponent scrub typhus vaccine candidates. Therefore, we cloned the 22-kDa gene from *O. tsutsugamushi* Karp into a DNA vaccine plasmid VR1012 (pKp22) and evaluated its protective efficacy in a lethal-challenge mouse model. To our surprise, the preliminary results suggested that pKp22 not only did not provide any protection but it also inhibited the protective effect of other scrub typhus vaccine candidates (Ching and colleagues, unpublished data). We do not understand these results, however we still believe that the phylogenetic relationships based on the sequences of 22-kDa genes are very valuable in selecting *O. tsutsugamushi* strains to generate components for vaccine purpose. We are in the process of determining 22-kDa sequences from additional strains of *O. tsutsugamushi*.

[*Competing interests statement:* The authors declare that they have no competing interests.]

REFERENCES

1. CHING, W.-M., H. WANG, C. EAMSILA, *et al.* 1998. Expression and refolding of truncated recombinant major outer membrane protein antigen (r56) of *Orientia tsutsugamushi* and its use in enzyme-linked immunosorbent assays. Clin. Diagn. Lab. Immunol. **5:** 519–526.
2. HICKMAN, C.J., C.K. STOVER, S.W. JOSEPH & E.V. OAKS. 1991. Molecular cloning and sequence analysis of a *Rickettsia tsutsugamushi* 22 kDa antigen containing B- and T-cell epitopes. Microb. Pathog. **11:** 19–31.
3. SEONG, S.Y., M.S. CHOI & I.S. KIM. 2001. *Orientia tsutsugamushi* infection: overview and immune responses. Microbes Infect. **3:** 11–21.